Chapter 2

Competition in banking

2.1 Introduction

In this chapter, we give an overview of the literature on the issue of competition in banking. But before discussing competition in the specific context of the banking sector, we briefly discuss competition in general. The word competition refers to the fact that firms (or banks) contend for patronage by customers or the public by attempting to supply or work at more favorable conditions than rivals. Firms can compete using various instruments, which can be distinguished based on the speed with which they can be changed (Tirole, 1988, p. 205). First, prices can often be altered in the short run. Second, cost structures and product characteristics can usually be adjusted only in the longer run. With respect to instruments that can be changed in the short run, we can consider competition in quantities as an alternative to price competition. In that case each bank chooses an output quantity and the price results from the inverse market demand function. However, in the real world we often observe banks competing in prices, i.e. interest rates, instead of quantities.

Market power is caused by a lack of competition, and determines a firm's ability to charge a price above marginal cost and thus earn a positive profit. This ability depends on the elasticity of demand facing the firm at the profit-maximizing quantity. If demand is highly elastic, a firm would lose a substantial amount of sales if it raised its price slightly. However, with inelastic demand, the firm loses fewer sales from the same price increase. A common measure for market power is given by the Lerner (1934) index or relative markup. This index expresses the

markup of price over (marginal) cost as a fraction of the price, that is

$$\text{Lerner index} = \frac{\text{price} - \text{cost}}{\text{price}}.$$

The greater the markup, the greater the Lerner index and the more market power a firm has. The above discussion indicates that the Lerner index is related to the elasticity of demand (Cowling and Waterson, 1976).

Now we turn to the discussion of the literature on competition in banking. First, we consider several approaches to banking that prevail in the theoretical banking literature. We discuss the main functions of banks, and different theoretical analyses of bank behavior. Then we look at competition in banking and discuss social welfare implications. Should individual banks' power be limited by having a large number of banks, as in the US, or is it desirable from a social point of view to have a highly concentrated banking sector, such as for example in the UK and in The Netherlands? Further, we discuss one of the approaches to banking, the Industrial Organization (IO) approach used in this thesis, in more detail. We focus on theoretical models that can be used to describe bank behavior in imperfectly competitive markets. Finally, we review some empirical evidence on the degree of competition in banking sectors in various countries.

This chapter should not be seen as a complete survey of the literature for this entire thesis. In each of the following chapters we will present a discussion of the literature relevant to the subject under consideration. Our main purpose here is to provide an introduction to the issue of competition in banking and to describe the Industrial Organization (IO) approach to banking.

2.2 Approaches to banking

The various approaches to banking that can be distinguished in the literature all focus either on one of the main functions of banks, or on a particular type of analysis of bank behavior. Below, we first discuss what it is that banks do. That is, why do banks exist? What are their main functions? Then, we turn to the question how individual banks perform these tasks, and give a short overview of theoretical explanations of bank conduct.

With respect to the inputs and outputs of banks, two different views can be taken. In the first view, a bank is seen as an ordinary service provider. In this view, which is sometimes referred to as the 'real resource approach' to banking (see Baltensperger, 1980), outputs are both loan and deposit services, and inputs should be interpreted as labor, fixed assets, materials, etc. In this interpretation, a bank is not that different from an ordinary firm offering services with more or less the same inputs. In the second view, which we consider throughout this thesis, banks are quite different from other firms. The focus here is in particular on commercial banks, which are characterized by the combination of the management of both loans and deposits. These banks are seen as financial *intermediaries*, whose outputs are loans and inputs are deposits (and, of course, the inputs mentioned above). Financial intermediaries reallocate 'the resources of economic units with surplus funds (savers or depositors) to economic units with funding needs (borrowers)' (Allen and Santomero, 2001, p. 273). That is, in this view the bank is seen as a middleman who transforms deposits into loans. We consider this intermediation or asset transformation role as the first (most important) function of banks.

This transformation function refers to different aspects of financial products. There may be many small depositors and few large lenders who wish to lend indivisible amounts. In that case, banks transform the denomination, i.e. unit size, of financial securities. Also, in this case small investors may not be able to diversify their portfolio, and the bank may be able to offer better risk-return features by quality transformation. This might also occur when there is asymmetric information and the bank is better able to process information than depositors are themselves. Further, transformation may concern the maturity of securities. Depositors are generally offered short maturities, whereas lenders often desire longer maturities.

We have argued that the most important function of banks is the transformation of financial securities. This raises the question why banks' customers do not transform securities themselves. The answer is given by transaction costs (including e.g. search costs and costs of monitoring) which may lead to economies of scale or scope. Such economies imply that the bank can more cheaply provide the transformation service (Gurley and Shaw, 1960; Benston and Smith, 1976). See also Santomero (1984) and Bhattacharya and Thakor (1993).

The second function of banks is to manage risk. This includes credit

risk, default risk, funding or withdrawal risk, interest rate risk, liquidity risk, and solvency risk. Banks must appraise the (credit) risk on bank loans in order to determine the risk-return characteristics of the loan. Default risk occurs when a borrower is not able to repay its loan plus interest payments. Funding or withdrawal risk refers to the risk that depositors withdraw their funds. The difference in maturity between loans and deposits discussed above induces interest rate risk. Liquidity risk refers to the difference in marketability of securities held and securities issued. Banks, as ordinary firms, may not be able to repay their creditors on time. Solvency risk emerges when the total value of a bank's liabilities exceeds the value of its assets. The latter two types of risk are not special to banks, but are relevant for any type of firm. However, in banking the consequences of these risks are far more serious, since the functioning of financial markets and financial stability is at stake (see section 2.3). For a discussion of the literature in this risk management approach to banking, see for example Baltensperger (1980), Santomero (1984), and Freixas and Rochet (1997, chapter 8).

The third function of banks is the processing of information on clients and the screening and monitoring of borrowers. Asymmetric information between banks and borrowers leads to two types of problems: adverse selection and moral hazard. Clients usually have more information on their characteristics and future prospects than banks have. Adverse selection refers to possible opportunistic behavior of such informed borrowers, who may benefit from contracting with an uninformed bank. That is, the bank cannot observe the borrower's (project's) quality, and precisely those borrowers with more risky (low quality) projects are generally eager for credit and willing to pay higher interest rates. This constitutes a problem to the uninformed bank. Moral hazard is opportunistic behavior of borrowers after they have obtained the loan, for example by not exerting sufficient effort to make their project a success. Thus, moral hazard refers in particular to unobserved behavior of borrowers. Banks generally invest in screening the different demands for loans they are confronted with before investing in new projects. This should inform them about the borrowers' characteristics. Also, banks invest in monitoring current projects, in order to limit the risk that the borrower does not invest the loan as it was originally agreed, or that he is unable to repay the loan plus interest. Because of the reusability of information, monitoring may lead to long-term relationships between a bank and a borrower. It may also lead to banks specializing in certain industries,

because information gathered about one firm or project may be useful for other projects in that same industry as well. (On the information processing function of banks, see also Santomero, 1984; Bhattacharya and Thakor, 1993; and Freixas and Rochet, 1997, chapters 2, 4, and 5 in particular).

Finally, banks offer access to a payment system, by which funds can easily be transferred from one bank account to another. Thus, they provide a public good whereas they are themselves private institutions. This provides a justification for public intervention in the banking sector, since access to a safe and efficient payment system must be ensured. This issue is becoming more and more important because of the globalization and internationalization of financial markets.

Next to these four main functions of banks and corresponding approaches to banking in the literature, several other approaches to banking can be distinguished that are related to different theoretical models or analyses of bank behavior. We discuss below asset choice or portfolio models; models of bank runs and bank regulation; credit rationing models; and models that fit into the IO approach to banking.

Portfolio or asset selection models determine the optimal asset (and liability) mix. Of course, this is closely related to risk management, the second main function of banks described above. In fact, several authors treat risk management and asset choice together as a single approach to banking, focusing on a broader interpretation of bank portfolio choice (e.g. Santomero, 1984; Baltensperger, 1980). However, asset selection models are not only related to risk management. Other aspects of banking may be involved as well. In most models of portfolio selection, either the asset market is taken to be perfectly competitive, or banks are assumed to have some degree of monopoly power in the asset market. The first type of models focuses on risk and return criteria in determining the asset mix, and is therefore related to the risk management approach. In the second type of models, however, banks maximize profits by choosing the optimal portfolio. This type is closely related to the IO approach to banking (discussed below). (See Santomero, 1984.)

Again elaborating on risk management, the withdrawal risk implies that banks are subject to runs, where many depositors withdraw their funds at the same time. If such a run occurs at several banks at the same time, it is called a banking panic. (For a discussion of banking panics and crises see Allen and Gale, 2000, chapter 9.) Evidently, bank runs and panics affect the safety and soundness of the banking sector

and of financial markets in general. In order to prevent runs and panics the government may introduce a lender of last resort facility, or deposit insurance. Other bank regulations include capital or reserve requirements, and portfolio restrictions. For a discussion of bank regulation see Bhattacharya and Thakor (1993) and Freixas and Rochet (1997, chapter 9). We discuss the issue of financial stability and the trade-off with competition in more detail in the next section.

Models of credit rationing explain why borrowers are sometimes refused credit, even though they are willing to accept all price and nonprice conditions of the contract. The contribution of Stiglitz and Weiss (1981) shows that adverse selection may lead to credit rationing. The reason is that an increase in the lending rate affects safer borrowers more than it does riskier borrowers, thus driving the safer borrowers out of the market. This could reduce the bank's expected return. The bank might therefore be unwilling to offer credit to a borrower, even if this borrower is willing to pay a higher rate. Similarly, with moral hazard an increase in the interest rate may increase the client's preference for risk, possibly reducing the bank's expected return. Moral hazard can therefore also lead to credit rationing. Thus, credit rationing models are closely related to the asymmetric information approach to banking. For overviews of the literature on credit rationing see Santomero (1984), Bhattacharya and Thakor (1993), and Freixas and Rochet (1997, chapter 5).

Finally, the Industrial Organization (IO) approach to banking applies models and ideas from the field of IO - the study of the functioning of markets - to banking. Although imperfections like transaction costs and asymmetric information explain why banks exist and can help to understand the conduct of banks, it is also important to investigate the functioning of the banking market itself. That is, in this study we are interested in the (strategic) behavior of individual actors on the banking market. This is not explained by the approaches to banking discussed above. In the IO approach, however, the focus is precisely on this functioning of banking markets and (imperfect) competition among banks. Empirical evidence indicates that banking sectors are often characterized by high concentration and at least some degree of market power (see section 2.5). Obviously, this may affect the functioning of banks, and thereby financial stability. The IO approach does not only study how banks compete, it also analyzes a wide range of other issues (e.g. related to regulation or asymmetric information) within the context of models of imperfect competition in banking. Since this approach to banking

provides the setting for this thesis, we will discuss it in more detail in section 2.4.

As the above discussion of various approaches to banking suggests, there is no comprehensive theory of banking which covers all important aspects of a bank. Instead, approaches and theories generally concentrate on some issues while neglecting others. This thesis is no exception to that rule. We focus on the IO approach to banking and study imperfect competition in banking. We also address some other issues, for example bank regulation, monetary transmission, credit rationing, and bank-firm relationships. Before discussing the IO approach to banking in more detail, we first turn to a brief discussion of competition and some social welfare issues with respect to competition in banking.

2.3 A competitive banking sector: Good or bad?

Deregulation, the decreasing influence of national borders, and increasing concentration due to mergers have made competition in the banking industry a major issue in the last decades. These changes attract much attention in the banking sector in particular. They do not only affect the individual banks' profitability and efficiency, but also the soundness and stability of the financial sector as a whole, and thereby the effectiveness of monetary policy instruments.

For a long time in the US competition in banking was thought to follow the general rule that competition enhances welfare. This result was simply copied from the traditional IO literature (see Tirole, 1988; Freixas and Rochet, 1997). The idea is that competition ensures efficiency, and thus maximizes social welfare. However, there is little reason to believe that such results will always carry over to the case of banks. Competition in banking is not necessarily a good thing from a social point of view. Banks differ from standard firms in several ways. For example, banks can be interpreted as intermediaries, as we argued above: they compete on both the output and the input market (that is, they are not price takers for inputs, as many ordinary firms are). This type of competition, first for inputs or deposits and then for outputs or loans, may cause inefficiency by itself. The intuition is that competition for deposits results in capacity constraints for the second stage of competition on the loan market. Because of these capacity constraints, the classic Bertrand result may not hold. In particular, when second-stage demand is relatively inelastic, banks have market power: prices are above the

Walrasian prices and there is excess stock. This implies inefficiency (Stahl, 1988). As a second example, note that increased competition may lead to excessive risk taking by banks. Banks use debt contracts with depositors, so they have limited liability which generally distorts incentives and implies excessive risk taking. This effect is stronger if competition is stronger and profits are lower. Therefore, a concentrated banking sector may actually be more efficient than a competitive one. (See Allen and Gale, 2000, chapter 8.)

In general, a less competitive banking sector implies higher profits. These profits can be used as a buffer in case of financial distress. In that sense, bank profits may support financial stability. Also, in a concentrated banking sector with a small number of large banks coordination and monitoring by the regulator will be easier than in a competitive banking industry with many small banks. Further, larger banks are less vulnerable because of their greater ability to diversify their portfolio and spread risks. This suggests that there is a trade-off between competition and stability in banking.

Historically, the US has had a very different view on competition in banking and thus a different development of the financial sector as compared to other industrialized countries (see Allen and Gale, 2000). There was a distrust of large financial institutions that have enormous power. Very early, this led to restrictions on banks to diversify or expand geographically, which ensured a large number of banks with limited power (e.g. the National Bank Acts of 1863 and 1864). It was argued that this would promote competition and lead to a competitive banking sector. Only after the severe banking crises of the early 1930s it became clear that the US view on competition in banking needed to be revised, and banks should be protected (to some degree) from competition in order to maintain financial stability.

Other countries, whether they have bank-oriented or market-oriented systems, have had highly concentrated banking sectors for a long time already. In the beginning of the twentieth century many industrialized countries' banking sectors became highly concentrated through a process of consolidation and the development of nationwide networks. Very often the governments of these countries stimulated these changes (Allen and Gale, 2000; Allen et al., 2001). Table 2.1 shows the number of commercial banks, the population per bank, and the three-firm concentration ratio (the percentage of total banking assets accounted for by the three largest banks) for five major industrial countries, and The

Country	Number of commercial banks	Population per bank	Three-firm concentration ratio[a]
United States	10,971	23,508	13.3
United Kingdom	491	118,328	29.1
France	425	135,365	63.6
Germany	330	245,379	89.5
The Netherlands	176	86,585	59.0
Japan	150	831,760	28.3

[a] Percent of total banking system assets accounted for by the largest three banks.

Table 2.1: Banking structure in various countries (1993).
Source: Barth et al. (1997, Table 3).

Netherlands. It illustrates the differences just discussed between the US and other industrialized countries such as the UK and The Netherlands. From the table, it becomes clear that the US has many more banks, less population per bank, and lower concentration in the banking sector.

Recently there has been important deregulation in many countries in particular with respect to 'the rules of competition' in banking, for example with regard to interest rates and bank activities. Two other pillars of bank regulation have been strengthened: prudential regulation has been developed further, and deposit insurance has been extended (see Allen et al., 2001).

Despite the arguments on the trade-off between stability and competition, recent research has shown that a more competitive banking sector has clear macroeconomic benefits. For the case of The Netherlands, Van Bergeijk at al. (1995) provide a rough estimate of the social cost of imperfect competition in the market for current account business credit, and argue that this cost is substantial (37-102 % of the actual interest payments). Smith (1998) presents a theoretical model to show that increased competition in banking may lead to a higher level of income and a reduction of the severity of business cycles. Intuitively, increased market power in banking leads to a higher cost of external funds (loans), but also of internal funds. The reason for this latter general equilibrium effect is that increased bank profits imply a higher return on bank liabilities, thereby increasing the opportunity cost of funds. Thus, all firms face higher costs of funding, which reduces the level of macroeconomic

activity. Smith argues that this effect is countercyclical. Concluding, increased bank competition (less market power) implies a higher level of macroeconomic activity, and less severe business cycles. Empirical evidence supporting the first hypothesis is provided by Cetorelli and Gambera (2001). Using data on 41 countries all over the world, they show that in general greater bank concentration has a negative effect on growth of the economy as a whole. This suggests that indeed, a more competitive banking sector is to be preferred. However, they also show that higher concentration in the banking sector tends to be correlated with higher growth of specific sectors of the economy, in particular those more in need of external finance. This would suggest an opposite result, at least for some sectors of the economy.

The discussion above shows that there is no clear-cut answer to this section's main question of whether competition is good or bad. On the one hand, a competitive banking sector is probably more vulnerable to liquidity shocks. Such shocks may lead to bank runs or panics when many depositors of a bank try to withdraw their money at the same time. On the other hand, competition may imply efficiency - although we have explained that this is not necessarily true under all circumstances.

2.4 The IO approach to banking

We argued above that in general it is not a good approach to simply copy some classical results from standard IO literature to explain bank behavior and its results. Despite this, IO-type of models can be fruitfully applied to banking when the specific character of banking activities is taken into account. This is exactly what the IO approach to banking does. It introduces particular characteristics of banks into IO models of competition to explain bank behavior and the development of banking markets. This methodological approach allows one to study the strategic nature of competition among banks as well as a wide range of other topics, e.g. financial stability, banking regulations, and the response of interest rates to a variety of exogenous or endogenous impulses (e.g. a policy rate change by the central bank).

In this section we present a survey of the literature in the IO approach to banking. We focus here on the theoretical analyses, and leave the discussion of the empirical evidence of imperfect competition to section 2.5. Some of the models listed below have already been mentioned in chapter 1, but will be discussed here in more detail. Originally, this

literature focused in particular on the so-called 'two-sided nature of the financial firm' (Santomero, 1984). This refers to the bank's main function as described in section 2.2 of reallocating resources from savers to borrowers. Financial intermediation by definition implies that banks compete for inputs (deposits) as well as for outputs (loans), and the volume of outputs is (at least to some degree) limited by the volume of inputs acquired. Thus, a bank acts as an intermediary and models of competition in banking are different from standard competition models. We start the survey below with a discussion of these models.

More recently, the IO approach has also turned to other aspects of banking. These are for a large part related to the various main functions and explanations of conduct of banks listed in section 2.2. That is, within the context of a model of imperfect competition among banks, these studies analyze issues that make banks special, i.e. different from ordinary firms. In general, this refers to characteristics of banks that either affect banks' strategic decisions and thus competition in banking (generally implying a departure from the perfect competition outcome), or stress the importance of market power in banking in particular. We discuss below IO models of banking concerned with asymmetric information, product differentiation, bank-firm relationships and switching costs, networks, competition versus stability, and monetary transmission.

In the remainder of this section, we sometimes refer to the effects of certain characteristics or events on interest rates. When we make mention of interest rates rising or falling, this should be interpreted as concerning lending rates, unless stated otherwise. In particular, for deposit rates we will generally have the opposite effect (circumstances that cause lending rates to rise, generally cause deposit rates to fall).

2.4.1 The 'two-sided nature' of the bank

Many early IO models of bank behavior are extensions of the models by Klein (1971) and Monti (1972). These models do not only determine the asset and liability mix but thereby also the size of the bank's portfolio, and they stress the type of competition in the industry. The models are described in surveys of banking by Baltensperger (1980) and Santomero (1984).

Klein (1971) described a single-period model of a representative profit-maximizing commercial bank that determines the optimal struc-

ture of its various types of assets and liabilities and the total size of the portfolio. The bank is assumed to hold three types of assets, i.e. loans, government securities, and cash. The liabilities of the bank consist of demand deposits, time deposits, and equity, where the latter is assumed to be exogenously given. The rate of return on government securities is exogenous. However, the bank is assumed to be a price setter in the other markets, i.e. the market for loans and the markets for both types of deposits. Thus the bank maximizes its profits, facing loan demand and deposit supply functions. The model by Monti (1972) is similar to Klein's (1971) model, but examines the implications of different objective functions such as deposit volume maximization. Therefore, the term 'Klein-Monti model' is often used to refer to adjusted versions of the models of Klein (1971) and Monti (1972).

In order to present a simple model to illustrate the main ideas behind the analyses of Klein (1971) and Monti (1972) we assume that there is only one type of deposits and one type of loans (see also Freixas and Rochet, 1997, pp. 57-58). There is a single, monopolistic bank who faces an upward sloping supply of deposits, given by $D\left(r_D\right)$, and a downward sloping demand for loans, $L\left(r_L\right)$. If the bank's volumes of deposits and loans are not equal, it can borrow or lend the difference on an interbank or money market at the exogenous rate i. Further, the bank is required to hold an exogenous fraction α of deposits as a non-interest bearing reserve at the central bank ($0 \leq \alpha < 1$). Finally, let $C(L, D)$ denote the costs of managing a volume L of loans and a volume D of deposits. For simplicity, let this function be linear, i.e. $C\left(D, L\right) = \gamma_L L + \gamma_D D$. Following Klein (1971), we assume that the bank maximizes its profits π by determining the optimal volumes of deposits and loans, D^* and L^*. Using the inverse deposit supply function $r_D\left(D\right)$ and the inverse loan demand function $r_L\left(L\right)$, the bank thus maximizes

$$\pi\left(L, D\right) = [r_L(L) - i]L + [i(1-\alpha) - r_D(D)]D - C(L, D).$$

Assuming that the function π is strictly concave, and r_L and r_D are continuously differentiable, the first-order conditions (FOCs) characterize the maximum and are given by

$$\begin{aligned} \frac{\partial \pi}{\partial L} &= r_L'\left(L\right) L + r_L\left(L\right) - i - \gamma_L = 0 \\ \frac{\partial \pi}{\partial D} &= -r_D'\left(D\right) D + i\left(1-\alpha\right) - r_D\left(D\right) - \gamma_D = 0. \end{aligned}$$

This shows that the optimal volume of loans L^* is determined by the first FOC only, and therefore it is independent of characteristics of the deposit side, i.e. the deposit supply function $D(r_D)$. Similarly, the optimal volume of deposits D^* is determined by the second FOC only and is independent of the loan side (see Langohr, 1982).

The original Klein-Monti model has been criticized by many authors, who suggest adjustments to the model in order to make its results more general. For example, Miller (1975) considers the introduction of different reserve requirements for the two types of deposits. Other extensions refer to the independence of the loan and deposit sides of the bank that is predicted by the model, or to the type of market structure. In chapter 1 we already mentioned that the model is based on a single, monopolistic bank but can easily be extended to the case of a finite number of identical banks that compete in quantities, i.e. Cournot competition (see Freixas and Rochet, 1997, pp. 59-61).

The most striking result of the original Klein-Monti model is that of independence of decisions on the loan side and decisions on the deposit side of the bank. That is, the optimal volume of loans (deposits) is independent of deposit (loan) market characteristics. Elyasiani et al. (1995) show how to test for asset-liability independence using individual bank panel data for the US in a dynamic econometric model. In the model, independence holds when loan demand and deposit supply only depend on (current, lagged, and expected future values of) the own rate and the federal funds rate. The empirical results suggest that independence does not prevail in general. With respect to theoretical considerations, the independence result does not hold if one relaxes the assumptions of the original model with respect to risk-neutrality, the slopes of demand and supply curves (see also Langohr, 1982), or the single decision period (Pringle, 1973). In each of these cases, demand characteristics as described by inverse loan demand $r_L(L)$ enter into the FOC for deposits, and vice versa. Evidently, the introduction of nonseparable management costs $C(L, D)$ implies that the loan side and the deposit side become interdependent. The introduction of liquidity and solvency constraints into the model also implies that the independence is lost (Van Loo, 1980). The same holds when bankruptcy risk and deposit insurance are added to the original model (Dermine, 1986).

Related to this literature, Sealey (1985) describes an alternative single-period banking model that also discusses the question of independence of the loan and deposit sides. This model focuses on financial

intermediaries owned by shareholders, and determines conditions under which independence, or portfolio separation, is in the interest of these shareholders. Sealey argues that because the existence of intermediaries implies the presence of incomplete markets, shareholder unanimity on optimal policies cannot be taken as given. Evidently, if shareholders disagree on the optimal policies, say on the optimal volumes of loans and deposits L^* and D^*, they cannot possibly agree on portfolio separation. Therefore, Sealey examines conditions for shareholder unanimity as well as conditions for independence to be in the interest of shareholders. Independence occurs in this model if and only if the shareholders agree that it will be beneficial, i.e. if both types of conditions are satisfied.

Another contribution to the literature on the Klein-Monti model is that by VanHoose (1988). He provides a partial equilibrium model based on the Klein-Monti model but with Cournot behavior of an endogenous number (denoted by n) of banks in a deposit market. He considers a deregulated market in which entry and exit are allowed, and combines the market supply curve and the market demand curve to determine the market clearing volume of deposits, under the free-entry condition that profits are zero. This model is used to discuss the possible effects of deregulation. The results show that the effects on the banks' deposit volumes of changes in other variables (such as interest rates or costs) change fundamentally if the number of bank rivals (i.e. n) changes. So, VanHoose concludes that deregulation - which affects n - may have a significant impact on bank behavior.

The models discussed above are all static models. Incorporating uncertainty or modeling lasting bank-client relationships in a model of bank behavior requires a dynamic model. We now summarize some theoretical models of bank behavior that consider two periods or two stages. Unfortunately, models of bank behavior with a longer time horizon appear to be limited in number and the 'dynamic' models in this section should be interpreted as models in which decisions taken in a specific order affect each other.

Prisman et al. (1986) introduce liquidity risk into the Klein-Monti model in a two-stage setting. In the first stage, the bank maximizes an expected profit function that consists of a part called 'normal profits' and a part that refers to the expected compensation to be paid when an infeasibility (i.e. a liquidity shortage) occurs in the second stage. Thus, a decision on the deposit volume must be taken in the first stage, prior to the value of the stochastic loan demand variable becomes known.

After observing this value, in case the liquidity constraint is violated in the second stage this infeasibility has to be resolved at a cost. That is, the bank has pricing power on the deposit market and uncertain loan demand may force the bank to attract a more expensive source of funds in the second stage. This implies again that independence of the loan side and the deposit side is lost.

Van Loo (1980) presents a simple two-period model based on lasting relationships between borrowers and lenders. These bank-customer relationships imply that the current volume of loans will affect future volumes of loans and deposits. So, current decisions of the profit-maximizing bank will affect its future profits. The model evidently results in interdependence of the loan side and the deposit side and in a higher loan volume and a lower loan rate in the first period to attract customers.

So far we have focused in this section on models where banks compete in quantities. It often seems more reasonable to think that banks compete by setting interest rates (prices) in practice, however. Fortunately, IO models of price competition can equally well be applied to banks. For example, Yanelle (1989) describes the results of a two-stage or two-period model of financial intermediation with Bertrand competition in both the loan and the deposit market. It is based on Stahl's (1988) double Bertrand model of (not necessarily financial) intermediation. The financial intermediation version of the model is described in some detail in Freixas and Rochet (1997). Surprisingly, the results differ from that of the standard single-period Bertrand model, i.e. from the Walrasian perfect competition outcomes. In the first stage of the model competition takes place on the input market. The quantity resulting from this stage determines the capacity for the intermediary in the second stage, when competition takes place on the output market. Thus, competition in the first market may imply an auction for the right to become a monopolist in the second market and the outcome of the game may therefore differ from the Walrasian outcome.

2.4.2 Asymmetric information

Asymmetric information is one of the main characteristics that make banks different from ordinary firms, and affect the strategic decisions of banks. We discussed in section 2.2 how adverse selection and moral hazard problems may lead to credit rationing. However, asymmetric in-

formation does not only affect the amount of credit a bank is willing to lend in equilibrium. It may also affect other strategic choices, such as entry and thereby market structure. That is, asymmetric information may represent a barrier to entry. This result was established by Dell'Ariccia et al. (1999) and Dell'Ariccia (2001). In a static model with Bertrand competition Dell'Ariccia et al. (1999) analyze the equilibria with two and with three banks, and show that the third bank does not want to enter. Intuitively, the potential entrant faces an adverse selection problem that is more severe than that of the incumbent banks. Note that entry is not endogenous in this setup. Dell'Ariccia (2001) solves this problem. He presents a dynamic model of spatial competition, in which entry is endogenous. He shows that even without fixed entry costs the equilibrium number of banks is finite because of the asymmetric information and 'learning by lending'. That is, asymmetric information implies a fixed cost (the precise value of which is determined endogenously), which is larger for new entrants than for incumbents - as in the simple static model of Dell'Ariccia et al. (1999). Although asymmetric information thus limits the number of banks entering the market, this does not automatically imply that interest rates are higher and competition is less aggressive. In fact, asymmetric information and 'learning by lending' induce banks to charge lower interest rates in order to compete for new borrowers (Dell'Ariccia, 2001). This implies that in the presence of (different degrees of) asymmetric information more concentrated banking markets may be characterized by lower interest rates.

In the presence of asymmetric information market structure may also affect the quality of borrowers that pass a bank's screening test. With adverse selection, competing banks often use imperfect screening tests to assess the riskiness of applicants' projects. Credit is offered only if the test indicates that the project is good. If the tests of different banks are independent, an increase in the number of banks systematically worsens the pool of applicants. Thus, the average creditworthiness of borrowers that pass a test decreases (Broecker, 1990). Also, more competition may make it more difficult for a given borrower to obtain funding. Riordan (1993) derives this result using a continuous quality signal. There is a cutoff value of this signal above which projects are funded. Riordan shows that a higher number of banks increases this cutoff value. This makes it more difficult for any borrower to obtain a loan, and can have a large negative effect on welfare.

Banks can alleviate these adverse selection problems to some extent

by using common filters, such as standardized credit scoring models or common databases. Increasing the number of banks does not affect the quality of applicants when banks only use common filters - in contrast to the case when filters are (to some degree) independent. In theory, the use of common filters can thus mitigate the adverse selection problem, even if their quality is slightly worse than that of the individual, independent screening methods (Shaffer, 1998). Furthermore, sharing information by creditors on their borrowers' credit records does not only reduce adverse selection problems, but it also works as a disciplinary device. This is related to the issues of moral hazard and monitoring. When information on defaults is shared, borrowers have an incentive to perform well, since in case of default it will be more difficult for them to obtain new funds at any bank. However, sharing more information (i.e. not only on defaults) may reduce these incentives (Padilla and Pagano, 2000). This problem can be solved by 'fine-tuning' the amount and type of information shared. In this way first-best incentives can be achieved. This may affect equilibrium interest rates as well.

The adverse selection problem and the resulting rationing of credit may alternatively be resolved by using collateral, instead of screening. Stiglitz and Weiss (1981) have shown that adverse selection and moral hazard may lead to equilibrium credit rationing. However, if banks use collateral as well as interest rates as strategic variables, there is no equilibrium credit rationing (Bester, 1985). That is, banks can use loan contracts as a self-selection mechanism. The borrower's willingness to post collateral renders information on his characteristics. This makes the adverse selection problem less severe and improves the allocation of credit. It also leads to lower interest rates. This may explain the widespread role of collateral; see also Chan and Thakor (1987) and Besanko and Thakor (1987). Replacing project screening by collateral may, however, not be desirable from a social welfare point of view. Banks may well be in the right position to judge the quality of borrowers' projects. Simply relying on collateral may be desirable for the banks themselves both because it mitigates the adverse selection problem and because of the disciplinary role of collateral (the borrower will exert high effort in order not to lose its collateral; i.e. the moral hazard problem is mitigated as well), and lead to lower interest rates, but screening may be more efficient from a social point of view. Too much reliance on collateral limits the banks incentives to screen projects and may lower welfare (Manove et al., 2001).

2.4.3 Product differentiation

Because banks in practice offer horizontally differentiated products or services, the spatial competition model (e.g. Salop, 1979) provides a convenient description of competition for banking. The spatial dimension of the model can either be interpreted as taste for a specific type of loan (or banking service in general) to incorporate the heterogeneous nature of products, or specifically as geographical location of banks or bank branches. It is well known that differentiated products imply at least some degree of market power. Thus, product differentiation in banking will generally imply a departure from the perfect competition outcome, and may affect market structure, location choice, and other strategic decisions of banks. Several banking studies apply oligopoly or monopolistic competition models based on horizontal differentiation or spatial competition as an instrument in order to study other specific aspects of banking (e.g. Chiappori et al., 1995 ; Matutes and Padilla, 1994; Dell'Ariccia, 2001; Schargrodsky and Sturzenegger, 2000). Here, we concentrate on models that focus on the differentiation issue and its effects on banks' strategic decisions itself.

Banks tend to have multiple branches and usually are active in several geographic markets. Each of these markets is characterized by horizontal differentiation; however, in general there are differences among banks in presence across markets. If banks are not allowed to price discriminate *across* markets, this may affect equilibrium interest rates. There may be price dispersion across banks *within* geographical markets, even if the banks have the same cost structures (Barros, 1999).

Next to interest rates, differentiation may also affect location choice. Consider the simple Hotelling model in which two banks choose locations in a linear city based on profit-maximization. Costs of monitoring borrowers can be interpreted as an increasing function of the distance between the bank and its borrowers. In this setup, two effects determine the optimal location (Wong and Chan, 1993). First, banks are drawn towards the endpoints of the linear city because of the competition effect: being further apart softens interest rate competition. Second, monitoring costs push the banks towards the middle of the city, because these costs are minimized if the bank is located in the middle of the part of the market it serves. The equilibrium locations therefore depend on the monitoring costs, and are generally different from the welfare-maximizing locations. This suggests that governments might want to

take bank locations into account when granting bank charters.

Now consider entry in geographical markets. Does the location of a market affect entry? Fuentelsaz and Gomez (2001) present an empirical analysis of the factors determining entry in local markets for the case of Spanish savings banks. When regulations limiting interregional banking were abolished, incumbent savings banks were confronted with the possibility of entering geographical markets that they were previously closed off from. Fuentelsaz and Gomez (2001) show that banks have a preference for entering nearby markets. Also, strategic interactions between incumbent banks and new entrants (such as entry deterrence by expansion of incumbents) appear to have affected the entry decisions of the Spanish savings banks.

Differentiation in banking may not only refer to horizontal (spatial) differentiation, but to vertical product differentiation as well. In particular, banks may specialize by offering certain additional products or services, such as phonebanking or remote access, that are preferred over the traditional services at least by some customers. Including both the geographical aspect and the specialization aspect leads to a model of multidimensional competition (Degryse, 1996; Bouckaert and Degryse, 1995). Specialization, e.g. offering remote access, allows a bank to steal clients from the competing bank since (some) clients like this facility. Also, specialization affects interest rate competition indirectly, since it affects the degree to which the two banks are substitutable from the clients' point of view. In particular, remote access implies that clients do not have to visit the bank's office, and thus reduces the degree to which the banks are horizontally differentiated.

2.4.4 Bank-firm relationships and switching costs

Very often we observe long-term bank-firm relationships. Banks develop close relationships with borrowers over time in order to reduce screening costs as well as monitoring costs. In that way, long-term relationships can mitigate asymmetric information problems. However, if a firm considers switching to another bank, that bank still needs to gather the information, the cost of which is commonly passed on to the client in the form of a fee. As a result, borrowers commonly face switching costs. Empirical evidence indicates that deposit markets are characterized by significant switching costs as well (Kim et al., 1999; Shy, 2002). This gives banks some degree of market power (see Klemperer, 1987; and

Sharpe, 1997).

Although bank-firm relationships help alleviate asymmetric information problems, there are (social) costs involved (see e.g. the survey on relationship banking by Boot, 2000). The costs refer to the soft-budget constraint problem and the hold-up problem. The soft-budget constraint problem refers to a bank's incentive to lend to a client in financial distress. In this situation, a bank that already lent to the firm may be inclined to provide additional funds even though a bank without such a relationship would not do so. The bank therefore cannot credibly commit to deny further credit. The hold-up problem concerns the informational advantage of banks. If a bank has gathered information on a firm, it can use this monopoly position to charge higher loan rates. The firm cannot avoid this by applying for credit at another bank, because the firm faces switching costs when turning to a different bank, as we argued above. This allows the firm's current bank to charge higher rates. The hold-up problem can be resolved by including a termination clause in long-term line of credit contracts. With such a contract, the borrower can continue the relationship if he chooses to do so, but in that case, the relationship is continued at pre-specified terms. This allows the bank to obtain some but limited bargaining power (Von Thadden, 1995). For additional references on relationship banking in general, see Boot (2000).

Bank-firm relationships are related to the issue of specialization discussed above. By engaging in a long-term relationship with borrowers, a bank can make itself more unique relative to other banks. First, a focus on relationship banking can be seen as specialization itself. Second, the relationship allows a bank to learn firm-specific information. The bank's knowledge of this information makes the bank unique from that particular firm's point of view. Thus, relationship banking can shield the bank from aggressive competition in the same way as (vertical) product differentiation can (see Boot and Thakor, 2000).

Increased competition in banking leads to lower rents and for that reason decreases the value of bank-firm relationships. However, increased competition from the credit market might lead banks to invest more in relationship banking, because this allows them to specialize and protects them against pure price competition. This suggests that increased credit market competition makes borrowers worse off (Petersen and Rajan, 1995). In the presence of deposit insurance and relationship banking, increased competition among banks may lead to more risk tak-

ing by the banks and in that way may hurt depositors as well (Besanko and Thakor, 1993).

2.4.5 Networks

Switching costs may dramatically change the way in which banks compete. A small fixed cost of switching between banks may imply that a market structure with a large number of independent competing banks will result in an equilibrium price at the monopoly level, whereas two banks that are allowed to have networks with several banking offices at different locations may charge a price that is close to that with perfect competition (Allen and Gale, 2000, chapter 8). This provides an argument for allowing branch banking, and confirms that market concentration is not always a good indicator of market power.

Next to branches or offices, networks in banking can also apply to technologies with network effects, such as Automatic Teller Machines (ATMs). ATM networks can be used as a competitive device in the deposit market. Again, this is related to specialization (see also the discussion of remote access and phonebanking above), and the same two effects described there are at work here. First, since depositors derive utility from ATMs, banks are encouraged to invest in their own network and to share their network with other banks. Second, however, sharing networks implies increased price competition. This effect lowers the bank's incentives to sign a sharing agreement. As a result, in equilibrium full compatibility of networks may not obtain (Matutes and Padilla, 1994). Economides and Salop (1992) provide a welfare analysis of networks, that can be applied to ATMs as well. However, in their setup, compatibility does not arise endogenously but is imposed exogenously. For a discussion of these models see also Shy (2001, chapter 8).

Saloner and Shepard (1995) provide an empirical analysis of the introduction of ATMs in the US in the 1970s, when ATMs were mainly installed at existing bank offices. They find evidence of network effects, which predict the value of the network to increase in the number of locations served. The empirical result that the adoption rate tends to increase with the number of offices (potential locations for ATMs) confirms this hypothesis. Also, adoption increases with deposit value, which confirms the production scale hypothesis (more users increase the value of the network). Thus, the strategic value of an additional ATM to a

bank, and therefore the incentive to install it, depends both on the size of the current network (including compatible ATMs of other banks) and on the number of clients of the bank that may use it.

2.4.6 Competition versus stability, and regulation

We have discussed several characteristics of banking, such as the two-sided nature, asymmetric information, differentiation, the presence of switching costs, and network effects, which in most cases imply less aggressive competition. That is, our discussion suggests that banking markets will be characterized by some degree of market power by nature. Now, we turn to some issues that make the competitiveness of the banking sector particularly important. That is, next to the standard social welfare effects of a loss of competition in an industry, market power in banking has some additional effects. From section 2.3, it is clear that one of these topics is financial stability.

Increased competition tends to destabilize the banking sector, because it leads to increased risk taking by banks (Broecker, 1990; Riordan, 1993; Besanko and Thakor, 1993). This issue has also attracted attention within the IO approach. For example, Matutes and Vives (2000) describe the effects of deposit insurance, focusing on competition for (heterogeneous) deposits and risk taking incentives. They show that the precise effects on welfare of bank regulation, such as asset restrictions, depend on the degree of competition in the deposit market as well as on the characteristics of the deposit insurance regime. Their results demonstrate that deposit competition in the presence of a deposit insurance regime can lead to excessive risk taking in banking.

In section 2.3 we already indicated that this view is not undisputed. We referred to the standard IO argument that competition increases welfare, which may hold in banking as well at least to some degree. We also argued that increased competition in banking may increase the level of income and reduce the severity of business cycles (Smith, 1998; Cetorelli and Gambera, 2001). The IO approach to banking has generated several more specific models that argue that competition in banking may enhance welfare. For example, Koskela and Stenbacka (2000) ask whether there is indeed a trade-off between competition and stability, as predicted above, when borrowers invest their loans in risky projects. They show that when banks condition interest rates on investment volumes, this is not necessarily the case. More precisely, in their model

increased competition for loans leads to lower interest rate and higher loan volumes, without increasing riskiness. In fact, lower lending rates decrease the probability of borrowers' bankruptcy. Schargrodsky and Sturzenegger (2000) present a related analysis of capital requirements. Such prudential measures are designed to promote financial stability, and are generally thought to lead to higher concentration in banking markets. In that sense, prudential measures would hinder competition. In a spatial competition model, Schargrodsky and Sturzenegger (2000) show that tighter capital requirements cause banks to choose a lower degree of specialization (in terms of horizontal product differentiation). This may lead to more aggressive competition. Thus, again, there need not be a trade-off between competition and stability.

2.4.7 Monetary transmission

Finally, consider the issue of monetary transmission, or more precisely the pass through of policy rate changes to bank lending rates. Clearly, since the nature of competition in banking affects interest rate setting, it influences the pass through and thereby the effectiveness of monetary policy measures. Swank (1994, section 2.4.3) gives a short overview of papers that address this issue. He argues that studies of monetary transmission in the context of imperfect competition models of banking have been limited because monetary policy aims are macroeconomic in nature (e.g. influencing growth or inflation). The microeconomic nature of imperfect competition models does not allow for a complete analysis of the effects of monetary policy. Nevertheless, a partial analysis of specific effects is possible and may yield interesting insights.

For example, Chiappori et al. (1995) analyze the pass through of policy rates to bank rates in the context of a spatial competition model. Under perfect competition, price equals marginal cost and monetary policy is fully effective in the sense that changes in the policy rate are fully passed on to bank rates. Also, with spatial competition among a limited number of banks and independent loan and deposit contracts, this result obtains. However, Chiappori et al. (1995) argue that deposit rate regulation makes it optimal for banks to offer tied sales, that is, to bundle loans and deposits. Intuitively, when deposit rates are regulated (at a level below the equilibrium deposit rate the banks would choose to offer) there is an incentive to lower credit rates in order to attract new deposits. With deposit rate regulation and tied sales contracts,

Chiappori et al. (1995) show that the pass through becomes incomplete. The induced change in bank rates will be less than proportional to a change in the policy rate in that case. Thus, deposit rate regulation may affect the pass through of policy rates to bank lending rates and therefore the effectiveness of monetary policy, via tied sales.

This type of analysis focuses on the effects of bank behavior on the transmission of monetary policy measures. There may also be a reverse effect, however. That is, monetary policy may affect bank behavior, and thereby the competitiveness of the banking sector. For example, in a model of implicit collusion among banks the interest rate setting rule used by the central bank may affect the incentive of the banks to collude over the business cycle (Bagliano et al., 2000).

2.5 Competition in banking: Empirical evidence

In this chapter, we discussed various approaches to banking, in particular the IO approach which is concerned with competition among banks, and we asked whether competition in banking is good or bad from a social welfare point of view. Before turning to the development of additional theories or the application of empirical methods to specific questions in the remainder of this thesis, we discuss here the empirical evidence on the actual degree of competition in banking. In chapter 1 as well as in section 2.3 we referred to concentration in banking markets and suggested that high concentration may point towards the absence of strong competition. However, we will argue below (section 2.5.1) that there may not be a simple one-to-one relationship between competition and concentration. Therefore, we discuss here some evidence about the actual degree of bank competitiveness (section 2.5.2).

2.5.1 Competition versus concentration

The competitiveness of a market is often thought to be related to market concentration. According to the Structure-Conduct-Performance (SCP) paradigm (see Gilbert, 1984), market structure influences the performance of firms in the sense that more concentrated markets facilitate collusive agreements, increase market power and therefore increase prices and profitability of firms. Evidence of such relationships may constitute an argument in favor of regulatory changes that aim to enhance competition in concentrated markets. The main hypothesis of the SCP

paradigm is that structure affects performance. This hypothesis can be tested empirically for the banking sector simply by regressing bank profits on a measure of concentration, e.g. the market share of the five largest banks, and some other explanatory variables (control variables that account for bank-specific and market specific characteristics). Evidence for the hypothesis is found when the coefficient corresponding to the concentration measure turns out to be positive and significant. The SCP paradigm contains two other (sub)hypotheses. First, structure is said to affect conduct. Second, conduct is thought to influence performance. Thus, as an alternative to regressing performance measures like profits on measures of market structure directly, the individual components of the paradigm can be tested by relating conduct to structure, or performance to conduct. In doing so, conduct can be measured by direct estimates of market power (see the next subsection).

Table 2.2 summarizes the main results of the survey by Gilbert (1984) and the more recent empirical studies of the SCP paradigm and the Relative-Efficiency hypothesis which we will discuss below. It also indicates the country under consideration, and whether the focus of the study is on loans, deposits, or both (where the latter may refer to the overall performance of banks). In general, the evidence (in terms of statistical significance) for the predictions of the SCP paradigm as well as for the Relative-Efficiency hypothesis is mixed.

The SCP paradigm evidently is not limited to studies of the banking industry. It is however remarkable that banking studies in particular usually estimate the paradigm directly without explicitly relating it to a formal model of bank behavior. An exception is the study by Calem and Carlino (1991), who incorporate 'conduct' explicitly by basing their empirical model on a theoretical model of bank behavior. This theoretical model is a liquidity management model similar to the Klein-Monti model (see section 2.4). A similar approach is taken by Hannan (1991b), who presents a theoretical banking model which is also based on the Klein-Monti model. He shows that the most important relationships predicted by the SCP paradigm can indeed be derived from this model.

Gilbert (1984) discusses the SCP paradigm and the results of 45 empirical studies of banking using this paradigm. He concludes that the results of those studies do not consistently accept or reject the hypothesis of a positive relationship between concentration and performance. Only about half of the studies he surveys find a significant positive coefficient. This mixed evidence also characterizes the limited recent empirical re-

Study	Country	Products	Result[a]
Gilbert (1984) - survey	US	(Various)	Mixed (half of the studies find evidence for SCP)
Berger and Hannan (1989, 1992)	US	Deposits	Evidence for SCP; no evidence for RE
Calem and Carlino (1991)	US	Deposits	No evidence for SCP
Hannan and Liang (1993)	US	Deposits	Evidence for SCP
Hannan (1991a)	US	Loans	Evidence for SCP; no evidence for RE
Goldberg and Rai (1996)	Europe	Both	No evidence for SCP; limited evidence for RE
Bikker and Groeneveld (2000)	Europe	Both	Evidence for SCP
Punt and Van Rooij (2001)	Europe	Both	Limited evidence for SCP; evidence for RE
Bikker and Haaf (2001)	US, Europe, and others	Both	Evidence for SCP
Lloyd-Williams and Molyneux (1994)	Spain	Both	Evidence for SCP; no evidence for RE

[a] SCP refers to the Structure-Conduct-Performance paradigm; RE refers to the Relative Efficiency hypothesis.

Table 2.2: Overview of SCP studies in banking.

search on the SCP paradigm, although some studies stress the positive relationships found (see for example Bikker and Groeneveld, 2000, and Bikker and Haaf, 2001, on European banking markets).

If a (significant) positive relationship between concentration and profits or performance is found, the SCP predicts that the effect runs via conduct. This is not necessarily true, though. An alternative explanation for such a positive relationship is provided by the Relative-Efficiency (RE) hypothesis (see Gilbert, 1984). This hypothesis states that efficient firms are able to earn relatively high profits because of lower costs, and thus increase in market share. This causes a high concentration of profitable firms. In this case regulatory changes that aim

at decreasing concentration are counterproductive, since it is desirable to have the most efficient firms active in a market. The empirical evidence on the question which hypothesis best explains the relationship between structure and performance is mixed.

Berger and Hannan (1989, 1992) argue that it is possible to test the SCP paradigm excluding the RE paradigm as an alternative. This should be interpreted as follows. When performance is measured by profits, the two paradigms are not mutually exclusive. Both non-competitive behavior and efficiency tend to increase profits, and the two effects cannot be distinguished. Berger and Hannan (1989, 1992) argue that when prices instead of profits are used to measure performance, the two hypotheses are mutually exclusive. More precisely, the SCP paradigm predicts that a more concentrated structure implies non-competitive behavior, which tends to raise prices (lending rates). Conversely, the RE hypothesis would, if anything, suggest lower prices because it predicts more concentrated markets to operate more efficiently. Note however that the two effects may be at work at the same time and the empirical test can therefore only examine which of the two explanations dominates. Berger and Hannan (1989, 1992) analyze US retail deposit rates, using prices (interest rates) instead of profits to measure performance. They find evidence that is consistent with the SCP hypothesis. Hannan (1991a) applies the same method to US loan rates and comes to the same conclusion.

Goldberg and Rai (1996) study the relationship between concentration and profitability by explicitly including measures of efficiency (such as X-inefficiency and scale-inefficiency) as explanatory variables in the econometric model. This allows them to test both the SCP hypothesis and the RE hypothesis directly. They estimate predictions of the two hypotheses for a sample of banks from 11 European countries and do not find any evidence of a relationship between concentration and performance. However, for countries with low concentration, they do find evidence of one of the predictions of the RE paradigm: for these countries, there exists a positive relationship between X-efficiency (that is, a superior production process) and profitability. For European banking markets, Punt and Van Rooij (2001) also find strongest support for the RE hypothesis, in particular with respect to X-inefficiency. This is contrasted by the findings of Lloyd-Williams and Molyneux (1994) for the case of Spain. Their evidence supports the SCP paradigm but not the RE hypothesis.

Hannan and Liang (1993) focus on the relationship between structure and conduct. They analyze bank-specific US deposit data in two steps. The first step is to estimate the market power (reflected by the conduct) of banks. The second step involves regressing those estimates on a measure of concentration and other explanatory variables. They find some evidence of a positive relationship between market power and concentration. Bikker and Groeneveld (2000) and Bikker and Haaf (2001) apply a similar approach to the case of Europe and find similar results.

The SCP and RE studies discussed above concentrate on cross sections of many industries, trying to identify general relationships between market structure, behavior, and performance. This involves several problems (see Bresnahan, 1989; Martin, 1993, chapter 17). First of all, this is a reduced form approach and the underlying mechanisms that govern strategic interaction and its results remain a black box. Second, the focus is on a limited number of observable variables, which may not be the only variables that matter. In particular, accounting measures like profits are simply taken as representative measures for economic performance. Furthermore, although empirical analysis of SCP relationships may yield some insight into the causes and effects of market power, and the relation between market structure and competitiveness, it does not allow one to assess the competitiveness of a given market or set of firms.

2.5.2 Competitiveness

The second, more recent, dominant method of empirical IO aims to infer market power or competitiveness directly. This so-called 'new empirical IO' is surveyed by Bresnahan (1989). In contrast to the SCP approach discussed above, which must almost by definition concern cross-section studies of many industries, the new empirical IO approach assesses conduct at the industry or even firm level. In general, it estimates behavioral equations that govern the price and quantity setting of the firm(s). One or more of the estimated parameters of these equations can directly be linked to analytical measures of competitiveness. This approach has been extensively applied to banking.

The new empirical IO includes various ways to test whether or not firms exert market power and to estimate the degree of competition in a market (see Martin, 1993, chapter 18). For empirical studies of banking, two methods dominate: first, the method of Panzar and Rosse

(1987), and second, the conjectural-variation method or its alternative specification generally referred to as the method of Bresnahan (1982) and Lau (1982). The method of Panzar and Rosse (1987) uses bank-specific data to estimate a reduced-form revenue equation. The explanatory variables in this equation include input prices. The sum of the estimated coefficients of input prices is called the Panzar-Rosse statistic H. This statistic can be interpreted as the sum of elasticities of gross revenue with respect to input prices, and it can be used to perform a test of imperfect competition. The second approach involves the estimation of the conjectural variation, the bank's anticipated response of its rivals to an output change. This response depends on the competitive behavior of the banks as a whole and thus the estimate can be used to infer the competitiveness of the sector (Iwata, 1974; Bresnahan, 1982; Lau, 1982). (For details, see chapter 3.)

Since there is not a one-to-one relationship between market structure, measured for example by concentration, and competitiveness (see section 2.5.1) one should be careful when interpreting the results of these empirical tests. For example, if the estimated value of the conjectural variation equals one, which in theory corresponds to a monopoly, this does not mean that there is a monopoly in the banking sector. Instead, there could be several banks active in the market and the result should be interpreted as banks behaving *as if* there was a monopoly. That is, for some reason or another these oligopolists are able to charge the monopoly price.

We now turn to the discussion of some empirical results, starting with the US. Table 2.3 presents an overview of the main results of the studies discussed below. It also indicates the country under consideration, the method used to estimate competitiveness, and whether the focus of the study is on loans, deposits, or both (where the latter may refer to the overall performance of banks). The empirical results are classified according to the degree of market power found, where 'no' refers to no (or hardly any) evidence of market power; 'yes' refers to evidence of a relatively high degree of market power; and 'limited' refers to either limited market power or mixed evidence.

Applications of the Bresnahan-Lau method to US banking data have not found evidence of strong market power. For example, Shaffer (1989) concludes that the loan market is more competitive than Cournot, and, in fact, he cannot reject perfect competition. Zardkoohi and Fraser (1998) analyze the effect of geographic deregulation in US loan markets.

Study	Country	Method[a]	Products	Evidence of market power[b]
Shaffer (1989)	US	BL	Loans	No
Zardkoohi and Fraser (1998)	US	BL	Loans	No
Shaffer and DiSalvo (1994)	Pennsylvania (US)	BL/PR	Loans	Limited
Cosimano and McDonald (1998)	US	Event studies	Both	Yes
Nathan and Neave (1989)	Canada	PR	Both	Limited
Shaffer (1993)	Canada	BL	Both	No
Molyneux et al. (1994)	Europe	PR	Both	Limited
Neven and Röller (1999)	Europe	BL	Loans	Yes
De Bandt and Davis (2000)	Europe	PR	Both	Yes
Bikker and Groeneveld (2000)	Europe	PR	Both	Limited
Suominen (1994)	Finland	BL	Both	Yes
Swank (1995)	Netherlands	BL	Both	Limited

[a] BL refers to the method of Bresnahan and Lau; PR refers to the Panzar-Rosse approach.

[b] The evidence of market power is classified as follows: 'no' refers to no (or hardly any) evidence of market power; 'yes' refers to evidence of a relatively high degree of market power; and 'limited' refers to either limited market power or mixed evidence (all in terms of statistical significance).

Table 2.3: Overview of empirical studies of market power in banking.

They show that in general deregulation did not have a significant effect on competition. In those states where the effect was significant, the sign was positive for some states and negative for others. A possible explanation for this result is that loan markets were already competitive before deregulation, which is confirmed by empirical testing. Shaffer and DiSalvo (1994) apply both the Bresnahan-Lau method and the Panzar-Rosse method to a duopoly banking market in Pennsylvania. The degree of competition turns out to be in between that in Cournot equilibrium and perfect competition.

Cosimano and McDonald (1998) use event studies in order to analyze the effects of elimination of reserve requirements. This allows them to examine whether market power exerted by banks in the US should be interpreted as market power of the aggregate industry relative to other financial institutions, or of individual banks within the banking sector. They argue that if the market power is on the interindustry level (i.e. competition between industries) then market structure within the banking industry is characterized by perfect competition and the benefits of the reserve requirement elimination will go to the consumers. If market power is on the intra-industry level (i.e. competition between banks) each individual bank exerts market power which it can use to receive the benefits of the elimination. The results of the event studies imply that market power in the US is in fact at the intra-industry level. Thus, their results provide evidence for market power in US banking.

Evidence for Canada, as for the US, generally indicates a relatively high degree of competition. For example, Nathan and Neave (1989) employ the Panzar-Rosse test statistic and find limited evidence of market power in the relatively concentrated Canadian financial markets. They argue that the characteristics of the Canadian financial system are in line with the predictions of contestability theory. According to this theory, the threat of potential competition may induce the banks to behave competitively. A later study of Canadian banks by Shaffer (1993) based on the Bresnahan-Lau method cannot reject the hypothesis of perfect competition. Shaffer further shows that regulatory and structural changes have not had much impact on the degree of competition.

For European banking markets, as for the US, the evidence is mixed, although European studies tend to find more evidence of market power. Molyneux et al. (1994) use the Panzar-Rosse method for major European banking markets, where 'major' refers to the largest banking markets in terms of total assets. The results indicate that banks in Germany, the

UK, France, and Spain behave as if under monopolistic competition (which is commonly used to refer to competition with differentiated products in this literature). For Italy, Molyneux et al. (1994) find a negative value for the Panzar-Rosse statistic, in which case the results are inconclusive. Neven and Röller (1999) use the Bresnahan-Lau approach and analyze the markets for mortgages and loans to corporate sectors in seven European countries. They find evidence of relatively strong market power, although the degree of market power decreases over time. De Bandt and Davis (2000) apply the Panzar-Rosse approach, and show that large banks are not fully competitive, but small banks seem to have even more market power. Bikker and Groeneveld (2000) also apply the Panzar-Rosse method to the European Union (EU) as a whole as well as to individual EU countries. They find a similar result of monopolistic competition in all but two EU countries. For the other two countries, Belgium and Greece, perfect competition cannot be rejected. Suominen (1994) applies a two-product version of the Bresnahan-Lau method to Finnish banks. He finds evidence of market power and, surprisingly, of a decrease in competition after deregulation. As an explanation Suominen (1994) suggests that the large increase in demand after deregulation may have facilitated collusion. For Dutch loan and deposit markets, Swank (1995) uses a dynamic version of the method of Bresnahan and Lau to show that the degree of competition is lower than in Cournot equilibrium. He also shows that the degree of oligopoly in the mortgage (loan) market has decreased, whereas in the market for savings deposits, it has increased.

Summarizing, the evidence on market power in banking is mixed. Some studies cannot reject the hypothesis of perfect competition (especially for the US) whereas others find strong evidence of market power. For Canada the degree of competition appears to be high, even though Canadian financial markets are relatively concentrated. For European countries the conclusion of monopolistic competition seems to dominate, although some studies cannot reject perfect competition. The empirical literature thus suggests that competition in banking is not perfect, but different from the monopoly or perfect collusion case in general. This illustrates that oligopoly models such as those discussed in section 2.4 are indeed relevant for the case of banks.

To conclude this section we take a quick glance at some recent developments in European banking which may have affected bank competitiveness. During the last decades there has been a consolidation trend in

the banking industries of many countries. For a discussion of the EMU case see Degryse and Bouckaert (2000). Berger et al. (1999) provide a survey of the literature on consolidation in banking. Evidently, mergers and takeovers influence the constitution of the banking sectors and therefore they usually affect how banks compete. Although a merger does not always result in an increase market power (see Berger et al., 1999; see also the mixed evidence on the SCP paradigm discussed above), the overall effect of the consolidation trend would probably be to weaken competitive forces. However, globalization and deregulation have affected banking market structures as well. For Europe, for example, we can think about e.g. the Second European Banking Directive (January 1993) which provided the legal framework for cross-border banking in the European Union; and the start of the EMU and the adoption of the single currency (January 1999). These developments have probably strengthened competitive forces. Despite this development, however, the empirical evidence discussed above suggests that European banking sectors are characterized by monopolistic competition. That is, for most of the evidence the two polar hypotheses of perfect competition and of monopoly or perfect collusion can be rejected, and banking sectors are characterized by monopolistic competition or oligopoly. Note however that the competitiveness may vary across submarkets. For example, banks may behave according to the perfect competition model in one submarket (as our evidence in chapter 3 suggests for the Dutch consumer credit market) whereas they do have market power in other submarkets (compare for example the evidence by Bikker and Groeneveld, 2000, who report evidence of monopolistic competition for Dutch banks as a whole).

2.6 Conclusion

In this chapter we reviewed the literature on competition in banking. We examined different approaches to banking, and argued that it is important to understand the functioning of banking markets, and the strategic behavior of individual banks. This issue is taken up by the IO approach to banking, which is the methodological focus of this thesis. Then we discussed bank competition from a social welfare point of view. The trade-off between competition and stability indicates that competition in banking is not necessarily a good thing. We presented a literature survey of the IO approach to banking. We described various theoretical

models based on imperfect competition and applied to banking. Next to bank behavior itself, and its consequences for e.g. banking market structure and social welfare, several other issues can be studied in the context of these models. We highlighted several of these issues, such as specialization, bank-firm relationships, networks, and bank regulation. Finally, we discussed empirical evidence which suggests that imperfect or monopolistic competition is indeed the relevant case for many banking markets in practice, in particular for case of the European ones.